

DOCUMENT RESUME

ED 401 812

HE 029 660

AUTHOR Gnagey, William J.; Potter, Kirsten I.
TITLE The Effects on Learning, Course Evaluation, and Team
Evaluation of Changing STAD Teams at Midterm.
PUB DATE [96]
NOTE 13p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; *College Students;
*Cooperative Learning; Course Content; *Course
Evaluation; Educational Psychology; *Grouping
(Instructional Purposes); Heterogeneous Grouping;
Higher Education; *Peer Evaluation; Peer Teaching;
Teamwork
IDENTIFIERS *Student Teams Achievement Divisions

ABSTRACT

This study is the seventh in a series of investigations on the effectiveness of the Student Teams Achievement Division (STAD) approach to cooperative learning among college students. The STAD approach features heterogeneous teams of four to six students who tutor each other on course materials and prepare each other for weekly quizzes. Students take the quizzes individually, but are awarded bonus points on the basis of their team's mean quiz score. A total of 196 undergraduates in four sections of an educational psychology course were divided into STAD teams of five or six students, with half of the students being placed in new teams halfway through the semester. T-tests were performed on both groups in regard to quizzes, examinations, course ratings, and team member performance appraisals. The study found that being assigned to a new team at midterm had little or no effect on learning the course subject matter, evaluating the course, or evaluating teammates. Copies of course and team member evaluation forms are included. (Contains 11 references.) (MDM)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED 401 812

THE EFFECTS ON LEARNING, COURSE EVALUATION, AND TEAM
EVALUATION OF CHANGING STAD TEAMS AT MIDTERM

William J. Gnagey & Kirsten I. Potter
Illinois State University

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☒ This document has been reproduced as
received from the person or organization
originating it.
- ☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

William J. Gnagey

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

THE EFFECTS ON LEARNING, COURSE EVALUATION, AND TEAM EVALUATION OF CHANGING STAD TEAMS AT MIDTERM

William J. Gnagey & Kirsten I. Potter
Illinois State University

Perspective

This experiment is the seventh in a series of investigations of the effectiveness for university students of Student Teams Achievement Divisions (STAD), one system of cooperative learning. These studies have been based on the research of Robert E. Slavin (1983, 1989, 1990, 1991) who focused primarily on STAD's use in public schools. This system features small heterogeneous teams of 4-6 members who tutor each other on the material in the course and prepare each other for weekly quizzes that measure chapter objectives. Students take the quizzes individually, but are awarded bonus points on the basis of the team's mean quiz score.

In our first experiment (Gnagey, 1988), four sections of sophomore educational psychology students (N=145) were divided into equivalent experimental and control groups. After a three week baseline period, both groups were taught in the STAD format for the rest of the semester. While students in the control group were awarded bonus points on the basis of the mean improvement of their teams, students in the experimental group received bonus points for their own individual improvement. Although no significant differences were found in the achievement of the two groups, the experimental classes rated the course as more effective than did the control classes.

In our second experiment (Gnagey & Ostrowski, 1991), it was hoped that the social loafing of some team members could be prevented by publicizing their individual contributions to their teams. Two classes of educational psychology students (N=75) randomly assigned to control and experimental sections were taught in a conventional lecture-discussion format for a four week baseline period. Both the experimental and control sections were taught using the STAD approach for the remainder of the semester. The control class received anonymous feedback concerning their team's quiz performance, whereas the experimental class received additional feedback making them privy to the scores of other members of their own teams. No significant differences were found between the mean quiz, midterm, and final exam scores for the two sections.

In our third experiment (Gnagey & Ostrowski, 1992), we tried to determine the differential effects of two ways of awarding bonus points within the STAD format. Two forty-student adolescent development classes and two forty-student educational psychology classes served as subjects in parallel experiments. During the first half of the semester, all classes were taught in the same STAD format in which bonus points were awarded to all members of the three teams with the highest mean chapter quiz scores. During the last half of the semester, students in one randomly chosen adolescent development class and one randomly chosen educational psychology class were awarded bonus points on the basis of their team's mean improvement points. The award structure of the other two classes remained the same. For each experiment,

the mean scores of the experimental and control groups were compared on seven quizzes, a final examination, and the University Course Rating Scale. In each experiment, the "mean improvement points" class outscored the "mean raw score" class on only one chapter quiz. It was concluded that the independent variable was not robust enough to materially influence either students' achievement or their evaluation of the course.

In our fourth experiment (Gnagey & Navarro, 1993), we compared the effects of two methods of constructing learning activities for classes being taught in the STAD format. One hundred fifty-four students in four sections of educational psychology served as subjects. During the first half of the semester, all sections were taught in the same STAD format during which the instructor authored and assigned activities designed to assist team members in preparing each other for the weekly chapter quizzes. During the second half of the semester, two randomly chosen experimental classes prepared and carried out their own activities, while the two control sections continued to use those prepared by the instructor. Comparisons of the experimental and control group means on five chapter quizzes, the final examination, and the University Course Rating Scale revealed no significant differences on any of these measures of achievement or the course evaluation. It was concluded that after one-half a semester's experience with instructor-constructed learning activities, student teams were able to plan and carry out some that were at least as effective as those constructed by their teacher.

In our fifth experiment (Gnagey & Navarro, 1994), we wanted to see if awarding bonus points for both mean quiz scores and mean improvement points would produce higher class achievement and course evaluation than either method by itself. One hundred fifty-four educational psychology students in four sections of educational psychology served as subjects in the experiment. At the beginning of the semester, all sections were taught in the usual STAD format with bonus points being awarded for correctly completing learning activities on Wednesdays, and for being in one of the four teams scoring highest on the weekly chapter quizzes. After Quiz 4, one section chosen at random began getting bonus awards for improvement scores while the other three sections continued being awarded points for high mean team quiz scores. After the midterm examination, two of these three teams were randomly assigned to other treatments. In one section, the two teams with the highest mean quiz scores and the two with the highest mean improvement scores were awarded bonus points. In the second, the four teams with the highest mean improvement points were rewarded. At the end of the semester, a final examination, and the University Course Rating Scale were administered to all sections. No significant differences were found on any of these comparisons save for two chapter quizzes. It was concluded that none of the experimental variations had a systematic effect on either learning or course evaluation.

In our sixth experiment (Gnagey and Denoyer, 1995) we compared process and product methods of awarding bonus points to STAD teams. One hundred sixty-five students in four sections of educational psychology served as subjects in the

experiment. They were assigned by the usual registration procedures employed by the University and were mostly sophomores taking their first course in a program designed to produce high school teachers.

For the first three weeks of the semester, all four classes were taught according to the usual STAD format. A General Linear Models Analysis indicated that there were no significant differences among the four sections on any of the first three quizzes. At this time, two experimental and two control groups were selected using a random number table. While the Wednesday teamwork and Friday quizzes remained the same for all sections, the teamwork for the experimental sections was graded as a project instead of being the basis for bonus points. Bonus points were awarded instead for mean quiz scores. In essence, the experimental subjects received bonus points for a learning product (quiz scores) while the control sections continued to be rewarded for a learning process (learning activities). T-tests were performed between the combined experimental and combined control group data for the remaining nine chapter quizzes, the midterm and final examinations, and the first and second administrations of the University Course Rating Scale (a course evaluation device) and the Team Member Performance Appraisal (a combined rating of the effectiveness of one's team members). Since no significant differences appeared in any of these analyses, it was concluded that teachers may award extrinsic reinforcers for either learning activities (the learning process) or quiz scores (the learning product) with equal results in the acquisition of subject matter, the evaluation of the course, and the attitudes developed among teammates.

Objectives

The objectives of the present experiment concern the differential effects of a one-team and a two-team approach to STAD. The following three hypotheses were tested:

1. Students who stay with the same STAD team for the entire semester, will learn more of the subject matter than students who change teams at midterm.
2. Students who stay with the same STAD team for the entire semester will rate the course as more effective than students who change teams at midterm.
3. Students who stay with the same STAD team for the entire semester will rate their teammates as more effective than students who change teams at midterm.

Methods and Techniques

One hundred sixty-nine students in four sections of educational psychology served as subjects in the experiment. They were assigned by the usual registration procedures employed by the University and were mostly sophomores taking their first course in a program designed to produce high school teachers.

During the second day of class, all students were assigned to one of eight five- or six-member teams by asking first the women and then the men to count off by eight's. Students with the same number became a team which met at a place designated by a diagram on the board. The first teamwork assignment was to learn each other's names and trade addresses and phone numbers.

For the first eight weeks of the semester, all four classes were taught according to the usual STAD format. Each Monday, the assigned chapter was introduced and appropriate material was presented by way of lectures, films or videos. Each Wednesday, heterogeneous teams of 4-6 members convened to work on projects which involved the application of the text material to practical situations. Bonus points were awarded for accurate work.

Each Friday after the corrected teamwork was returned, the class was divided into two activities. At the beginning of the hour, the instructor held a question and answer session in which students could ask for clarification of difficult concepts in the chapter. When there were no more questions, a 20-item multiple choice chapter quiz was administered to all students individually. At the end of the Friday class, the appropriate chapter assignment for the following week was written on the chalkboard.

During class on the following Monday before the new chapter was introduced, all quizzes were returned, and a list of "grades so far" was posted by student social security numbers. Students were encouraged to write rebuttals, by Wednesday, to any of the items they missed but felt they should have credit for. The STAD cycle then began for the new chapter.

At midterm, a 60-item multiple choice examination was administered over the first six chapters. The University Course Rating Scale (UCRS) was also administered at this time as was the Team Member Performance Appraisal instrument. (See Appendix 1).

The UCRS is composed of twelve items on which students are asked to evaluate all aspects of the course: clarity of objectives, projects and papers, textbook and other assigned readings, in-class activities, quizzes and exams, feedback, interestingness, instruction, grading procedures, grading fairness, amount learned, expected level of performance (standards), and group activities.

In a previous study of 158 students, (Gnagey & Ostrowski, 1992), factor analysis of the UCRS (see appendix) using varimax rotation revealed three principal factors (See Table 1.). The first accounted for 31% of the variance and loaded heavily on items D (in-class activities), G (interestingness), H (instruction); J (amount learned), K (expected level of performance), and (group work). The second principal factor accounted for 29% of the variance and loaded heavily on items A (clarity of objectives), B (out of class papers and projects), C (text), E (quizzes and exams), and I

(fairness of grades). The third factor accounted for only 11% of the variance and loaded on item F (feedback).

The test-retest coefficient of reliability for the UCRS total score was .81 using 48 students in similar classes with one administration at midterm and the other eight weeks later during the finals.

A Team Member Performance Appraisal (TMPA) rating scale (see Appendix 2) was administered requiring all students to rate themselves and the other team members of their teams on the following criteria: a) team meeting attendance, b) contribution of ideas, c) completion of assignments, d) promoting positive feelings among team members, e) encouraging the expression of other team members. The ratings were done using a 5-point scale for each criterion: one indicating "almost never" and five indicating "almost always".

After the midterm examination at the end of the eighth week, a random number table was used to designate two sections as experimental groups while the other two became control groups which continued to be taught in the same manner as before. T-tests were used to compare the combined experimental and combined control group scores on all dependent variables before the experimental treatment began (See Table 2.) No significant differences appeared in any of the nine comparisons. Students in the experimental sections were then placed in new teams using the same counting-off procedure that was employed at the beginning of the semester.

At the end of the semester a 60-item multiple choice examination was administered to all four classes covering only those chapters assigned since midterm. The UCRS and TMPA were also administered for the second time.

Results and Conclusions

T-tests were performed between the combined experimental and combined control groups for all measurements (13 quizzes, the mid-term and final examinations, and the first and second administrations of the UCRS and the TMPA. Only one significant difference appeared in all of these analyses. (See Table 3). The control classes scored higher on a Calculation quiz than did the experimental sections.

Educational Importance

Under the conditions of this experiment, it appears that being assigned to a new team at midterm has little or no effect on learning that subject matter, evaluating the course, or evaluating one's teammates. Since forming new teams involves extra time and record keeping for the instructor, it is probably not an efficient approach to STAD in college classes.

References

- Gnagey, W.J. (1988). College level STAD: Group Vs Individual Reward. Paper read at Midwestern Educational Research Association meeting, Chicago.
- Gnagey, W.J. & Denoyer, J. W. (1995). Process vs. Product Methods of Making STAD Team Awards. Paper read at Midwestern Educational Research Association Meeting, Chicago.
- Gnagey, W.J. & Navarro, N. Y (1993). A Comparison of Student-constructed vs. Teacher- constructed Learning Activities in Phase Two of the STAD Cycle. Paper read at the Midwestern Educational Research Association meeting, Chicago.
- Gnagey, W.J. & Navarro, N.Y. (1994). The Effects of Changing the Criteria For STAD Team Awards. Paper read at Midwestern Educational Research Association Meeting, Chicago
- Gnagey, W.J. & Ostrowski, K. (1991). STAD in College. Paper read at Midwestern Educational Research Association meeting, Chicago.
- Gnagey, W. J. & Ostrowski, K. (1992). The Effects of Changing the Criteria for STAD Team Awards at Midterm: Two Parallel Experiments. Paper read at Midwestern Educational Research Association meeting, Chicago.
- Sherman, L.W. (1986). Cooperative vs. Competitive Educational Psychology Classrooms: A Comparative Study. Teaching and Teacher Education, 2, 283-295.
- Slavin, R.E. (1983). When Does Cooperative Learning Increase Student Achievement? Psychological Bulletin, 94, 429-445.
- Slavin, R.E. (1989). Research on Cooperative Learning: Consensus and Controversy. Educational Leadership, 47, 52-54.
- Slavin, R.E. (1990). Cooperative Learning: Theory, Research and Practice. New Jersey: Prentice Hall.
- Slavin, R.E. (1991). Synthesis of Research on Cooperative Learning. Educational Leadership, 48, 71-82.

Table 1

Rotated Factor Pattern for UCRS2

<u>Item</u>	<u>Factor 1</u>	<u>Factor 2</u>	<u>Factor 3</u>
A	0.07836	0.69193	0.12490
B	0.36320	0.58192	-0.30593
C	0.19670	0.58731	-0.40395
D	0.77268	0.25255	0.06187
E	0.27006	0.71172	0.12786
F	0.10361	0.23254	0.79150
G	0.60411	0.45104	-0.18749
H	0.69302	0.41617	-0.00718
I	0.03367	0.76926	0.19525
J	0.64069	0.39456	-0.32949
K	0.73521	-0.09374	-0.00112
L	0.69043	0.07623	0.10489

Percent of Variance Explained by each Factor

<u>Factor 1</u>	<u>Factor 2</u>	<u>Factor 3</u>
31.31526	29.25415	11.11909

Table 2

Pre-Treatment T-Test Comparison of All Dependent Variables

<u>Treatment</u>	<u>N*</u>	<u>Var</u>	<u>$\bar{X}$</u>	<u>SD</u>	<u>t</u>	<u>DF</u>	<u>P</u>
C	83	Q1	13.8	3.4			
E	86	Q1	14.0	3.3	0.42	167	0.68
C	83	Q2	14.8	2.7			
E	84	Q2	15.4	2.6	1.52	165	0.13
C	82	Q3	29.7	5.9			
E	86	Q3	29.2	6.2	0.52	166	0.60
C	81	Q4	13.4	2.5			
E	85	Q4	12.9	2.6	1.03	164	0.30
C	79	Q5	13.6	2.3			
E	84	Q5	13.9	2.7	0.56	161	0.57
C	81	MID	41.7	7.4			
E	84	MID	40.4	7.3	1.11	163	0.27
C	75	UCRSI	45.8	7.9			
E	72	UCRSI	46.5	6.8	0.59	145	0.56
C	79	TMPAI	23.0	1.9			
E	83	TMPAI	22.4	3.3	1.57	132	0.12
C	83	UCRSIL	3.4	1.4			
E	86	UCRSIL	3.2	1.7	1.11	167	0.27

* The variation in N is due to absences and withdrawals.

Table 3

Post-Treatment T-Test Comparison of All Dependent Variables

Treatment	N*	Var	$\bar{X}$	SD	t	DF	P
C	83	Q7	13.1	3.7			
E	86	Q7	12.9	4.1	0.37	167	0.71
C	83	Q8	13.1	3.8			
E	86	Q8	12.9	4.5	0.31	167	0.76
C	83	Q9	12.3	4.1			
E	86	Q9	12.7	4.4	0.71	167	0.48
C	83	Q10	12.0	4.3			
E	86	Q10	11.5	5.1	0.62	167	0.53
C	83	Q11	13.7	3.4			
E	86	Q11	13.0	4.9	1.12	153	0.26
C	83	Q14	12.0	3.8			
E	86	Q14	11.4	4.9	0.92	160	0.36
C	83	QCAL	13.1	5.6			
E	86	QCAL	10.6	5.9	2.85	167	0.01
C	83	FIN	34.3	10.2			
E	86	FIN	34.7	12.5	0.26	167	0.80
C	83	UCRS2	40.5	16.4			
E	86	UCRS2	39.5	17.2	0.41	167	0.68
C	83	TMPA2	21.2	5.9			
E	86	TMPA2	20.2	7.3	0.97	167	0.33
C	83	UCRS2L	3.1	1.6			
E	86	UCRS2L	3.1	1.6	0.02	167	0.99

* The variation in N is due to absences and withdrawals.

TEAM MEMBER PERFORMANCE APPRAISAL

Directions: In the space provided, list in alphabetical order by last name, all members of your team including yourself. Rate the contribution of each member on all of the five criteria listed below. Put an asterisk (*) in front of your name.

- A. Attends team meetings.
- B. Contributes helpful ideas to team discussions.
- C. Completes team assignments on time.
- D. Promotes positive feelings among team members.
- E. Encourages other team members to express themselves.

Use the following 5-point scale to rate each criterion for each team member.

- 5 = almost always
- 4 = often
- 3 = sometimes
- 2 = seldom
- 1 = almost never

Team Member Names	Ratings of criteria					Sum
	A	B	C	D	E	
1						
2						
3						
4						
5						
6						
7						
8						

UNIVERSITY COURSE RATING SCALE

- _____ A. The objectives of this course were clear to me. I knew what I was expected to learn. (1) hardly ever, (2) occasionally, (3) sometimes, (4) frequently, (5) almost always.
- _____ B. The assigned, out-of-class projects, papers, etc., helped me fulfill the course objectives. They assisted me in mastering the appropriate materials and skills. (1) hardly ever, (2) occasionally, (3) sometimes, (4) frequently, (5) almost always.
- _____ C. The textbook and/or other assigned readings helped me fulfill the course objectives. They assisted me in mastering the appropriate materials and skills. (1) hardly ever, (2) occasionally, (3) sometimes, (4) frequently, (5) almost always.
- _____ D. The in-class activities planned by the instructor helped me fulfill the course objectives. The experiences I had in class assisted me in mastering the course materials and skills. (1) hardly ever, (2) occasionally, (3) sometimes, (4) frequently, (5) almost always.
- _____ E. The measurement devices used in this course were accurate indicators of the extent to which I was fulfilling the objectives. They allowed me to show what I learned. (1) hardly ever, (2) occasionally, (3) sometimes, (4) frequently, (5) almost always.
- _____ F. The feedback in this course was adequate. During the semester, I knew how well I was doing. (1) hardly ever, (2) occasionally, (3) sometimes, (4) frequently, (5) almost always.
- _____ G. My interest in the course was kept high enough to motivate me to do good work. I was able to apply myself. (1) hardly ever, (2) occasionally, (3) sometimes, (4) frequently, (5) almost always.
- _____ H. In my opinion, the teaching of this course was: (1) poor, (2) fair, (3) average, (4) good, (5) excellent.
- _____ I. The grades I received for this course were fair. They were what I deserved for what I learned. (1) hardly ever, (2) occasionally, (3) sometimes, (4) frequently, (5) almost always.
- _____ J. Compared with other courses I have taken at this institution, I feel that I learned: (1) much less than in most courses, (2) less than most, (3) about an average amount, (4) more than in most, (5) much more than in most.
- _____ K. Compared with other courses I have taken in this institution, I feel that the level of performance expected of me was: (1) much lower than most, (2) lower than most, (3) about the same as most, (4) higher than most, (5) much higher than most.
- _____ L. The group activities carried out in class helped me fulfill the course objectives. My team helped me master the course materials and skills. (1) hardly ever, (2) occasionally, (3) sometimes, (4) frequently, (5) almost always.